Fully 3D Monte Carlo Reconstruction in SPECT: Proof of Concept and is it Worthwhile ?

Irène Buvat, Delphine Lazaro, Vincent Breton

***Abstract*—In Single Photon Emission Computed Tomography (SPECT) with parallel hole collimation, image reconstruction is usually performed as a set of 2D analytical or iterative reconstructions. This approach ignores the 3D nature of scatter and detector response function that affect the detected signal. To deal with the 3D nature of the image formation process, iterative reconstruction can be used by considering a 3D projector modeling the 3D spread of photons. This has already been studied using approximate analytical models for the 3D projector. In this paper, we investigate the value of using accurate Monte Carlo simulations to determine the 3D projector including all physical effects affecting the imaging process (attenuation, scatter, camera point spread function) used in a fully 3D Monte Carlo (F3DMC) reconstruction approach. Given the 3D projector, the reconstruction problem is solved using the maximum likelihood expectation maximization (MLEM) algorithm. To validate the concept, 2 small datasets were simulated and 4 reconstruction strategies were compared: filtered backprojection, MLEM without attenuation correction, MLEM including all corrections with approximate analytical models (MLEMC) and F3DMC. Results suggest that F3DMC multiplies imaging sensitivity by about 10^3, increases signal-to-noise ratio by 25 to 70% compared to MLEMC and improves spatial resolution. The practical feasibility of the approach on real data sets is discussed.**

***Index Terms*— Image reconstruction, Monte Carlo simulations, SPECT.**

I. Introduction

In Single Photon Emission Computed Tomography (SPECT) with a parallel hole collimator, image reconstruction is usually performed as a set of 2D analytic or iterative reconstructions. Each 2D reconstruction considers the 2D sinogram associated with a transaxial slice, assuming that photons emitted in a transaxial slice are detected within a single column in each planar 2D projection. This is an approximation because of the spatial response of the gamma camera and because of scatter. In the real world, photons emitted in a transaxial slice are not only detected in the projection column facing the slice, but also in neighboring columns. To better satisfy the 2D hypothesis, scattered photons can be subtracted from the projections before reconstruction, but this reduces the sensitivity of the imaging process. The finite point spread function can also be partially compensated for by filtering the projections before reconstruction. An alternative approach is to deal with the real 3D nature of the problem, using an iterative reconstruction algorithm involving a 3D projector which models the spread of photons in 3D due to scatter and spatial response function. This has been proposed using approximate analytical models (e.g., [1-4]). The concept of using Monte Carlo simulations to estimate the 3D projector has also been proposed early [5,6], but not applied in fully 3D at that time, due to impractical storage and computation time. As computer science is evolving fast, fully 3D Monte Carlo (F3DMC) reconstruction in SPECT might become feasible soon. Our aim was thus to determine the gain to be expected from F3DMC. F3DMC uses Monte Carlo simulations to estimate the 3D projector including all physical effects (attenuation, scatter, camera point spread function) affecting the imaging process. The reconstruction problem using this projector is solved with a maximum likelihood expectation maximization (MLEM) approach. Using small simulated datasets, we compared F3DMC with other 2D and 3D reconstruction methods.

I. Buvat is with the U494 INSERM, CHU Pitié-Salpêtrière, 75 634 Paris Cedex 13, France (phone: 33 1 53 82 84 19; fax: 33 1 53 82 84 48; e-mail: buvat@imed.jussieu.fr).

D. Lazaro and V. Breton are with the Laboratoire de Physique Corpusculaire, CNRS/IN2P3, Université de Clermont-Ferrand, France (e-mail: lazaro@clermont.in2p3.fr, breton@clermont.in2p3.fr).

II. Theory

A. Formulation of the fully 3D reconstruction problem

A discrete expression of the SPECT tomographic reconstruction problem can be as follows:

$$p = \mathbf{R}\, f,$$

where p is a column vector with PxN^2 elements – assuming P projections of N x N pixels are acquired –, f is a column vector of N^3 elements – assuming N transaxial slices N x N are to be estimated – and $\mathbf{R}$ is a (PN^2,N^3) matrix corresponding to the fully 3D projector. An element r_{ij} of matrix $\mathbf{R}$ corresponds to the probability that a photon emitted in voxel j is detected in projection pixel i. Because the problem is huge (typically, $\mathbf{R}$ is a 262,144 x 262,144 matrix if N=P=64), it is not addressed in its full dimensionality. Instead, it is usually factorized as a set of N independent reconstruction problems involving projections with PxN

elements, objects with NxN elements, and a (PN,N^2) projector. Assuming that the dimensionality of the problem might become tractable soon, our aim was to determine the gain to be expected by solving the fully 3D problem compared to alternative reconstruction approaches.

B. *Modeling* ***R***

To solve the fully 3D reconstruction problem, the first step is to model the projector **R**. Monte Carlo simulations were used to estimate **R**, as they currently offer the best modeling of all physical effects involved in SPECT. The **R** elements depend on the geometry and attenuation properties of the object as well as on the characteristics of the imaging system (fixed for a given imaging protocol). Assuming the attenuating properties of the object under investigation are know, from a CT scan for instance, **R** is estimated by running a Monte Carlo simulation that considers a uniform activity distributed over the attenuating medium. For each detected event, the couple (j,i) is stored, where j represents the emission voxel and i represents the detection pixel. From all detected events, the r_{ij} element of matrix **R** is deduced as the ratio of the number of events emitted in voxel j and detected in pixel i over the number of events emitted in voxel j. **R** can be calculated for any energy window. For storage savings, the dimension of **R** is not (PN^2,N^3), but only (PN^2,M), where M is the number of voxels belonging to the attenuating medium. Indeed, it is fair to assume that only voxels belonging to the attenuating medium contribute to the observed projections.

C. *Reconstructing the image volume*

Given the projector, the inverse problem p = **R** f can be solved using classical iterative algorithms. Because the measured projections follow Poisson statistics in SPECT, we used the MLEM algorithm. The result is the f column vector with M elements, representing the activity distribution within the attenuating medium.

III. Material and Methods

A. *Simulated phantoms*

Two phantoms were considered. “Phantom 1” was a 10 cm diameter and 10 cm height water filled cylinder, including a 2 cm diameter sphere filled with water. The sphere was centered in the cylinder and a Tc-99m activity of 24 MBq/ml was set in the sphere, while no activity was introduced in the cylinder (Fig. 1a). “Phantom 2” was a 10cm x 10 cm x 10cm tank consisting of 3 layers of different attenuation media along the X direction (Fig. 1b) corresponding to air, water and bone. This phantom included a set of line and point sources (Fig. 1b) with no activity in the background. The relative activity concentrations were 6, 8 and 10 in the “1”, “2”, “3” point sources and 10, 15 and 20 in the X, Y and Z line sources.

For both phantoms, the volumes to be reconstructed were sampled on a 10 x 10 x 10 voxel grid (1 cm^3 voxels).

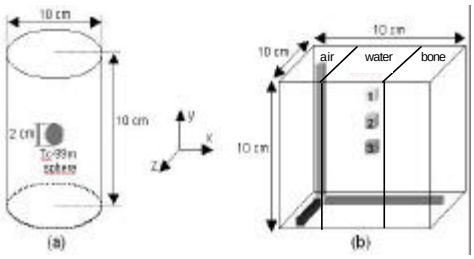

Fig. 1. Water cylinder (a) and tank (b) phantoms.

B. *Monte Carlo simulations*

Monte Carlo simulations were performed using the code GATE (http://www-iphe.unil.ch/~PET/research/gate/) that has been recently validated for various configurations in SPECT. Photon transport within the collimator was simulated using the Monte Carlo approach, making the simulations inefficient, but ensuring an accurate model of the imaging system and sensitivity. No variance reduction techniques were used.

C. *SPECT configurations*

For each phantom, a SPECT acquisition of 64 projections 10 x 10 was simulated with a 12 cm radius of rotation. The gamma camera characteristics were chosen to mimic those of the AXIS (Philips). About 100 million photons were generated and 105,649 were detected between 126 and 154 keV for phantom 1, whereas about 162 million photons were generated and 195,914 were detected for phantom 2.

D. *Projector calculation*

For phantom 1, a simulation with a uniform Tc-99m activity distribution within the water cylinder was performed. About 2 billion events were simulated and about 2 million events were detected between 126 and 154 keV. The simulation took about 97 CPU hours on a biprocessor Pentium III 1GHz machine. Similarly, for phantom 2, 2 billion photons corresponding to a uniform Tc-99m activity distribution within the non-uniform attenuating medium were simulated. For each phantom, the matrix **R** was deduced as explained before.

E. *Image reconstruction*

To assess the value of F3DMC, data collected in the 126-154 keV energy window were reconstructed using 4 methods:

1) filtered backprojection (Ramp filter, ν_c=0.5 $pixel^{-1}$): FBP.

2) MLEM without attenuation correction, 30 iterations: MLEM.

3) MLEM with attenuation correction (attenuation modeled in the projector), Jaszczak scatter correction [7], and 3D correction for depth-dependent spatial resolution using an analytical model [8], 60 iterations: MLEMC.

4) F3DMC including corrections for scatter, attenuation and finite spatial resolution with 30 iterations of MLEM:

F3DMC.

FBP and MLEM are 2D reconstruction methods, while MLEMC and F3DMC are fully 3D reconstruction methods. MLEMC uses "analytical" models for scatter and spatial response corrections, while F3DMC uses Monte Carlo models.

F. Image assessment

The reconstructed images were assessed in terms of:

- Reconstruction efficiency, defined as the number of events in the reconstructed volume divided by the number of simulated events,
- Signal-to-noise ratio (SNR). For each phantom, 20 noisy replicates of the projections were obtained. Each replicate was reconstructed using the 4 reconstruction methods. SNR was defined as the mean number of counts within an ROI (sphere for phantom 1 and 4 hottest pixels of the Z line for phantom 2) averaged over the 20 replicates of reconstructed images, divided by the standard deviation of that mean.
- For phantom 1, number of "mislocated" events, defined as the total activity detected outside the 8 voxels containing the sphere divided by the total reconstructed activity.
- For phantom 2, spatial resolution: in-plane and axial spatial resolutions were assessed by drawing a x-direction profile and a z-direction profile through the hottest point source (number 3) and estimating the FWHM of these profiles.
- For phantom 2, relative quantitation. For each line source, an average activity value was determined by averaging the value of the 4 hottest pixels. The ratios between the average activity measured in the Z and Y line sources (theoretical value = 20:15 = 1.33), and between the average activity of the Z and X line sources (theoretical value was 20:10 = 2) were considered.

IV. Results

For phantom 1, reconstruction efficiency, SNR and percentages of mislocated events for the 4 reconstruction methods are summarized in Table 1. For phantom 2, reconstruction efficiency, SNR, in-plane and axial spatial resolutions and quantitation indices are given in Table 2.

TABLE I
IMAGE ASSESSMENT FOR PHANTOM 1

Reconstruction method	Reconstruction efficiency	SNR	Percentage of mislocated events
FBP	$1.2\ 10^{-3}$	182	60
MLEM	$1.1\ 10^{-3}$	168	48.2
MLEMC	$2.1\ 10^{-3}$	182	3.9
F3DMC	$9.8\ 10^{-1}$	315	2.7

TABLE II
IMAGE ASSESSMENT FOR PHANTOM 2

Reconstruction method	Reconstruction efficiency	SNR	Spatial resolution (in plane/axial)	Relative quantitation Z:Y / Z:X
Ideal	1	-	-	1.33 / 2
FBP	$1.1\ 10^{-3}$	326	1.75 / 1.58	0.84 / 2.07
MLEM	$1.2\ 10^{-3}$	408	1.48 / 1.76	0.71 / 1.53
MLEMC	$1.9\ 10^{-3}$	61	1.16 / 1.23	1.26 / 1.90
F3DMC	$8.9\ 10^{-1}$	78	1.09 / 1.02	1.02 / 1.99

V. Discussion

A. Implementing the F3DMC approach

Implementing F3DMC requires:

1) The knowledge of the attenuating properties of the object to be reconstructed. This could be obtained from a CT of the patient, ideally performed on a hybrid SPECT/CT machine.

2) An accurate Monte Carlo simulation code to calculate the projector. The relevance of the projector obviously impacts the value of the reconstructed data. The amount of work currently dedicated to Monte Carlo simulations makes it possible to say confidently that accurate Monte Carlo simulations will be more and more widespread and fast in the future.

Although presented here for parallel hole SPECT, the method could be easily generalized to fan-beam or cone-beam SPECT reconstruction, and to fully 3D PET reconstruction (with even greater storage issues than for SPECT though).

B. Value of the F3DMC approach

Tables 1 and 2 suggest there is much to be gained by accurately modeling the 3D projector involved in iterative reconstruction, especially in terms of image spatial resolution, imaging sensitivity, and signal-to-noise ratio. As expected [8], the 3D approaches MLEMC and F3DMC significantly improve image quality and quantitation compared to the 2D reconstruction approaches (FBP and MLEM without correction). When the projector is unbiased, F3DMC should actually perfectly correct for scatter, attenuation and spatial resolution loss, unlike the 3D reconstruction approach MLEMC that uses an approximate scatter correction and an accurate analytical 3D spatial response correction. MLEMC was expected to work very well on phantom 1 with homogeneous scattering medium. Actually, on this example, the main advantage of F3DMC is the reconstruction of a much higher number of counts than MLEMC. This is not only because scattered photons are included in the reconstructed image while they are disregarded in the Jaszczak correction, but mostly because the count loss due to the collimator was modeled in the projector. This high reconstruction efficiency should result in a large SNR improvement. The limited SNR increase we observed compared to that expected from the sensitivity assessment is due to the noise in the projector. Simulating more counts when calculating the projector results in a greater SNR increase (results not shown).

The results obtained with phantom 2 confirmed those obtained for phantom 1 and illustrate the value of F3DMC for a non-uniform attenuating medium. As for phantom 1,

reconstruction efficiency was close to 1. SNR was better for F3DMC than for MLEMC but unlike for phantom 1, it was much lower for the 3D reconstruction approaches than for the 2D reconstructions without correction. This might be because of the non-uniform and non symmetrical attenuating medium of phantom 2, making attenuation correction challenging.

As suggested by the number of mislocated events in phantom 1, in-plane and axial spatial resolutions as assessed for phantom 2 were significantly improved by the 3D reconstruction approaches. In addition, F3DMC better restored high spatial resolution than MLEMC. This might be because F3DMC as implemented did not include any approximation. Its only theoretical limit lied in the uncertainties associated with the elements of the projector **R**, which resulted from the limited number of events used to estimate **R**.

Finally, the quantitative accuracy indices considered on phantom 2 did not consistently improve for 3D reconstruction compared to 2D (e.g., the Z:X ratio was more accurate for FBP than MLEMC). A more thorough assessment of the F3DMC and MLEMC merits regarding quantification is in progress.

The simulator used to create our data was identical to that used to calculate the projector, yielding the best achievable results. The actual value of F3DMC in practical configurations will be better determined by thoroughly studying the impact of the statistical noise present in **R** and by testing the robustness of the approach with respect to errors in defining the attenuating medium (e.g., from a CT) and in modeling the imaging process by the Monte Carlo code.

C. Numerical feasibility of the F3DMC approach

Given the huge size of the projector involved in the inverse problem to be solved, one can question the numerical feasibility of the method, from the points of view of storage, numerical stability, and computation time.

Assuming that 64 projections 64 x 64 are acquired to reconstruct a 64 x 64 x 64 volume, the projector would include 64^6 elements, i.e. would need 512 gigabytes for storage in double precision. Although this might appear prohibitive, efficient storage can make it tractable. As an example, with a generic compression algorithm (Lempel-Ziv), compression rates of 72% and 91% were obtained for the projectors used in the case of “phantom 1” and “phantom 2” respectively.

Considering the numerical stability of the method, the calculation of each projection elements would involve M additions (M is the number of voxels belonging to the attenuating medium), each addition term resulting from a multiplication. One can show that the overall error is less than 2^{-52} M X, where X is the maximum value in a projection bin. For $M = 256^3$ and $X = 10^5$, this error would remain less than 10^{-11}, suggesting that round-off errors will not be an issue.

Considering the computation time, two components must be considered. First, a Monte Carlo simulation has to be run for each dataset corresponding to a specific acquisition protocol and attenuation configuration (i.e., for each patient). The simulation duration will depend on the acceleration techniques available for the simulation code, on whether the photon transport through the collimator is simulated by Monte Carlo methods, and on the number of counts simulated to get a robust estimate of the projector. Several days of CPU are currently needed for simulating the projector corresponding to a patient acquisition with GATE, but acceleration techniques are currently developed to achieve computation time less than 1 day CPU. Once the projector is calculated, reconstruction time will depend on the iterative reconstruction algorithm and on the number of disk access and compression/decompression operations needed to read the projector. For the configurations presented in the paper, reconstruction time was about 2 min for 30 MLEM iterations on a Sun Sparc 20 workstation. Time for reconstructing 64 projections 64 x 64 would thus be about 8 hours. Using OSEM instead of MLEM will make it possible to reduce this time by a factor 16 at least. Adding the time required for disk access and compression/decompression operations suggests that overnight reconstruction might be perfectly realistic.

VI. Conclusion

Fully 3D Monte Carlo reconstruction appears worthwhile and might soon become practically feasible. This customized reconstruction approach makes use of the specific attenuation properties of each patient and of the specificity of the imaging system. Preliminary results suggest that it yields significant image improvement in terms of image signal-to-noise ratio, spatial resolution, and imaging sensitivity.

Acknowledgments

The authors thank the members of the OpenGATE collaboration and especially Dr. Christian Morel (Institute of High Energy Physics, University of Lausanne) who initiated the GATE project.

References

[1] Beekman FJ, Kamphuis C, Viergever MA 1996 Improved SPECT quantitation using fully three-dimensional iterative spatially variant scatter response compensation IEEE Trans. Med. Imaging 15 491-499

[2] Beekman FJ, den Harder JM, Viergever MA, van Rijk PP 1997 SPECT scatter modelling in non-uniform attenuating objects Phys. Med. Biol. 42 1133-1142

[3] Beekman FJ, de Jong HWAM, Slijpen ETP 1999 Efficient SPECT scatter calculation in non-uniform media using correlated Monte Carlo simulation Phys. Med. Biol. 44 N183-N192

[4] Bai C, Zeng GL, Gullberg GT 2000 A slice-by-slice blurring model and kernel evaluation using the Klein-Nishina formula for 3D scatter compensation in parallel and converging beam SPECT Phys. Med. Biol. 45 1275-1307

[5] Floyd CE, Jaszczak RJ, Coleman RE 1985 Inverse Monte Carlo: a unified reconstruction algorithm for SPECT IEEE Trans. Nucl. Sci. 32 779-785

[6] Floyd CE, Jaszczak RJ, Greer KL, Coleman RE 1986 Inverse Monte Carlo as a unified reconstruction algorithm for ECT J. Nucl. Med. 27 1577-1585

[7] Jaszczak RJ, Greer KL, Floyd CE, Harris CC, Coleman RE 1984 Improved SPECT quantification using compensation for scattered photons J. Nucl. Med. 25 893-900

[8] Tsui BMW, Frey EC, Zhao X, Lalush DS, Johnston RE, McCartney WH 1994 The importance of accurate 3D compensation methods for quantitative SPECT Phys. Med. Biol. 39 509-530